Luigi Spezia

Biomathematics & Statistics Scotland

Aberdeen

BAYESIAN VARIABLE SELECTION IN MARKOV MIXTURE MODELS

Based on joint works with...

Roberta Paroli (Università Cattolica SC, Milano)

Mark Brewer (Biomathematics & Statistics Scotland)

Susan Cooksley (The James Hutton Institute)

Christian Birkel (University of Aberdeen)

Overview

1) Markov mixture models:

Hidden Markov models

Markov switching autoregressive models

State-space models with regime-switching

Markov mixture transition distribution models

Mixed Hidden Markov models

Spatial hidden Markov models

. . .

2) Variable selection methods:

Stochastic Search Variable Selection

Kuo and Mallick's method

Gibbs Variable Selection

Metropolized Kuo-Mallick

3) Simulation results: Non-homogeneous hidden Markov model

Markov switching autoregressive models + covariates

4) Three applications: Bernoulli non-homogeneous hidden Markov model

Non-homogeneous Markov switching autoregressive models + covariates

Markov mixture models

Hidden Markov models

Markov switching autoregressive models

State-space models with regime-switching

Markov mixture transition distribution models

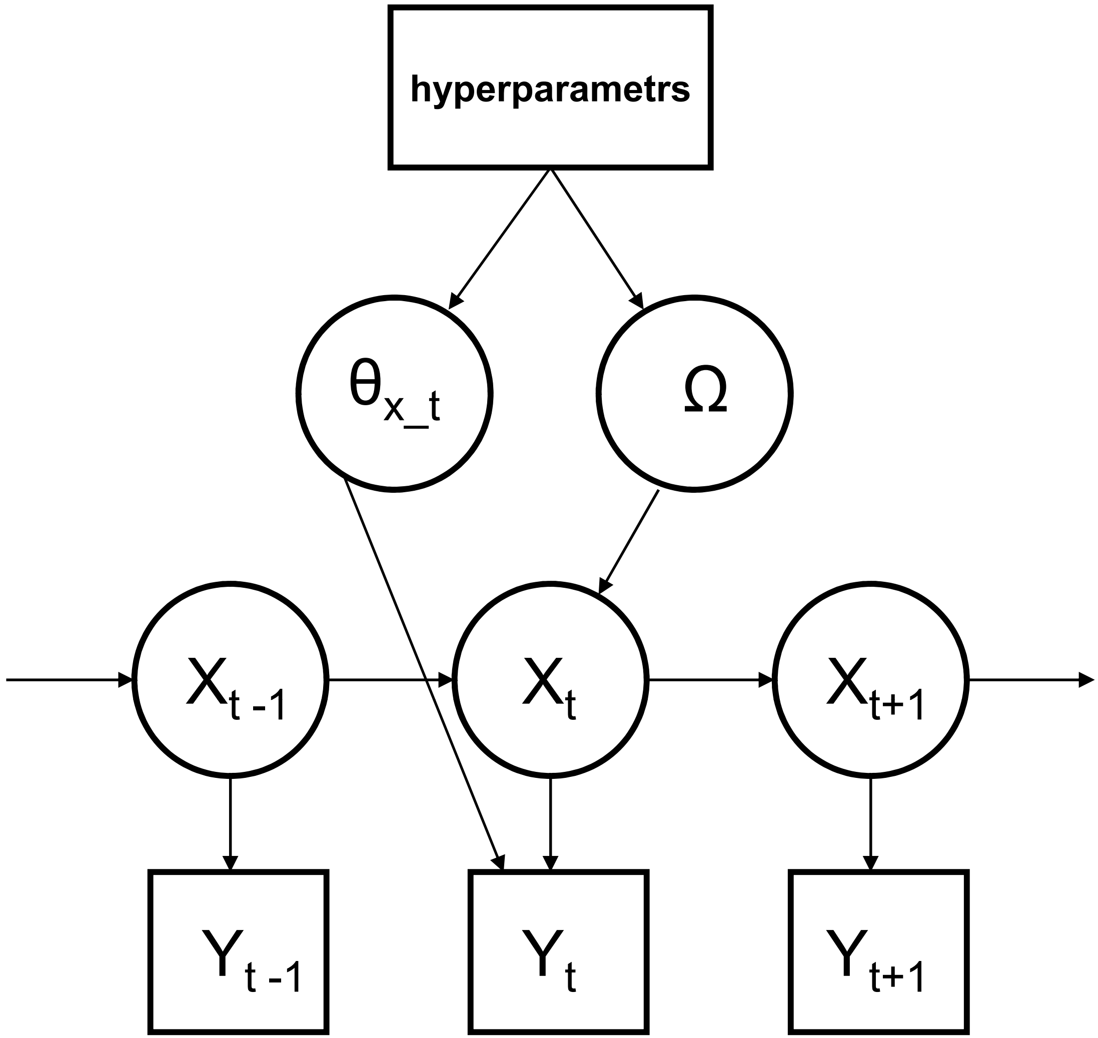
Mixed Hidden Markov models

Spatial hidden Markov models

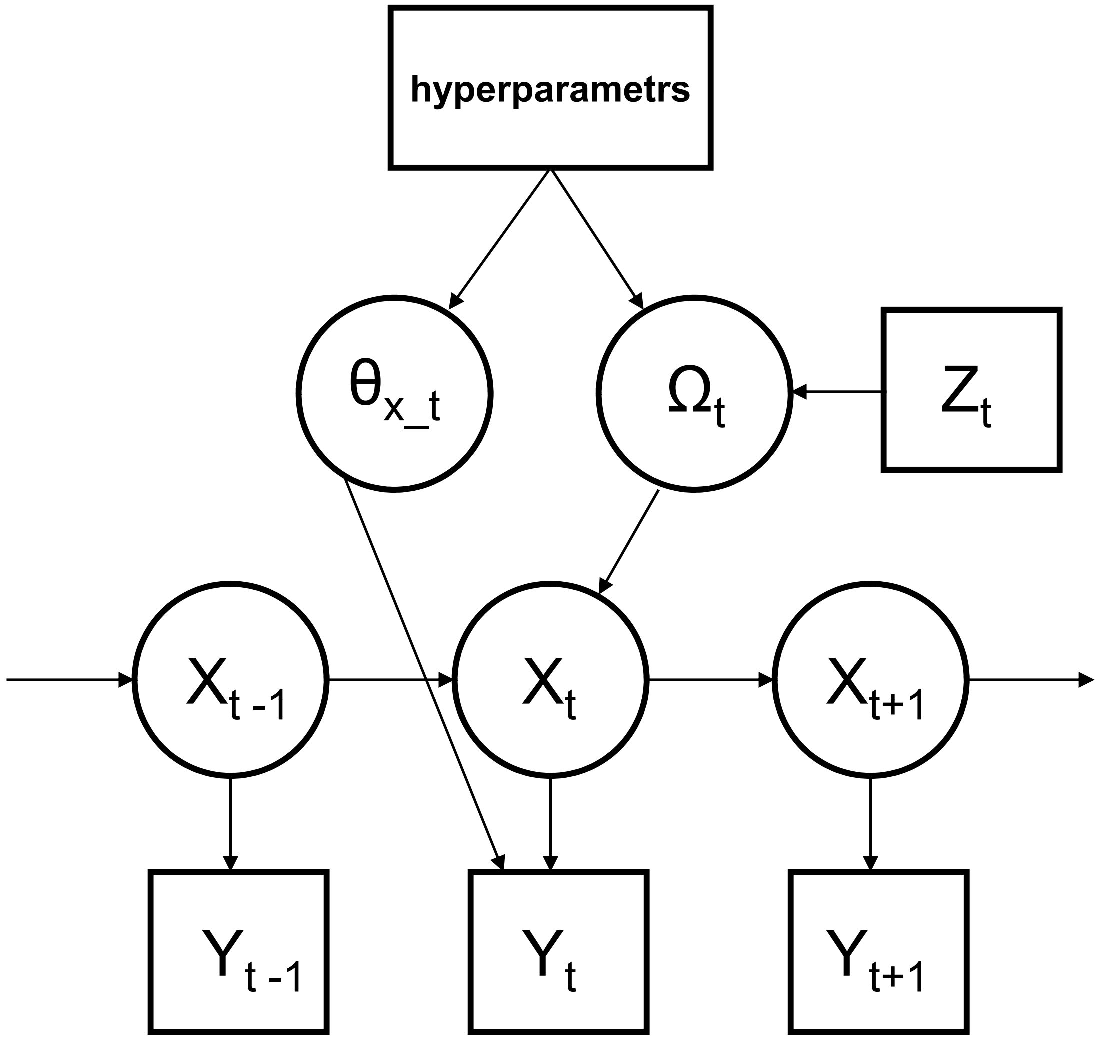
$$y_t \sim \sum_{j=1}^{m} \omega_j \, p(y_t | \theta_j) \qquad \sum_{j=1}^{m} \omega_j = 1$$

$$y_t \sim \sum_{j=1}^{m} \omega_{i,j} \, p(y_t | \theta_j) \qquad \sum_{j=1}^{m} \omega_{i,j} = 1$$

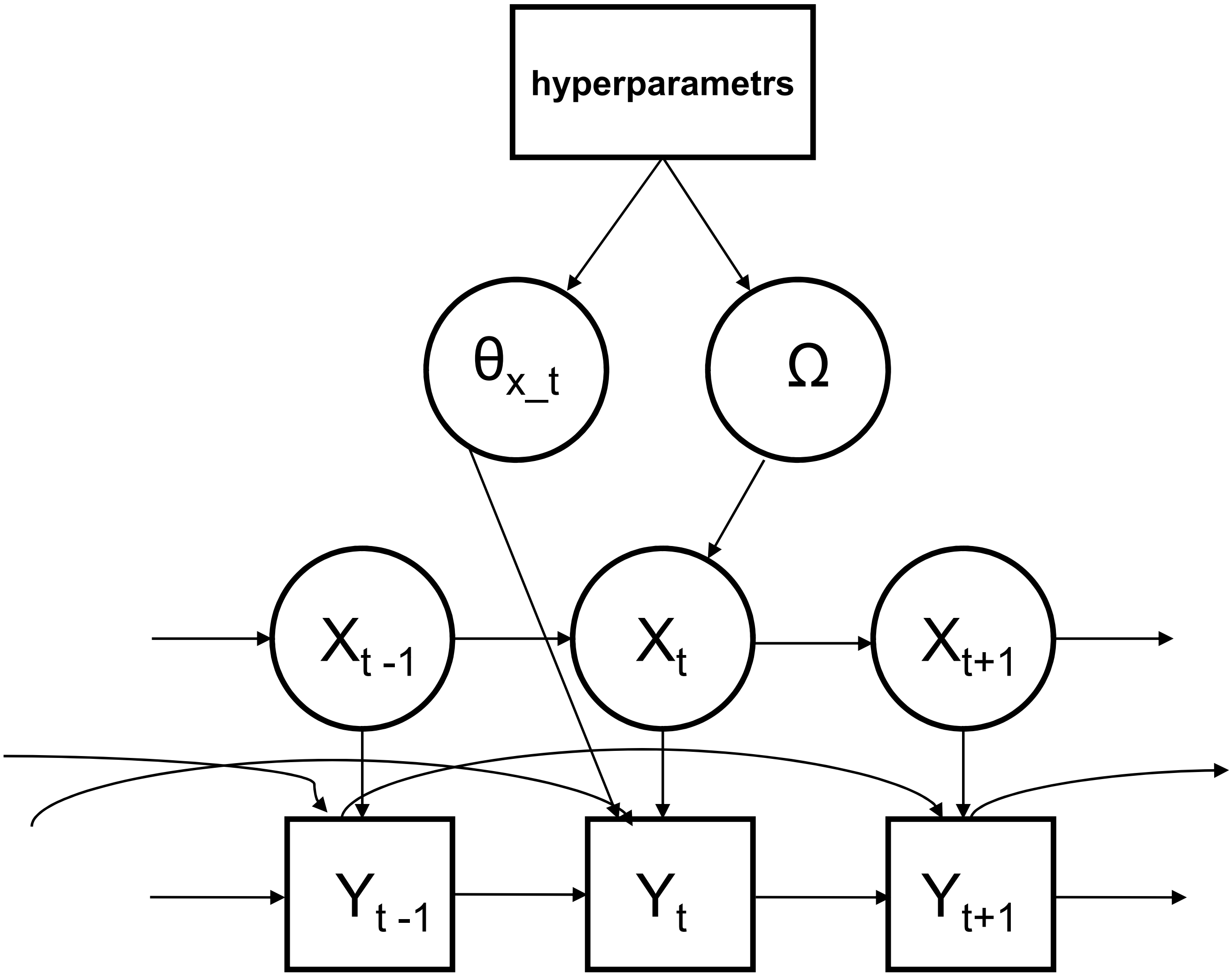
Hidden Markov models



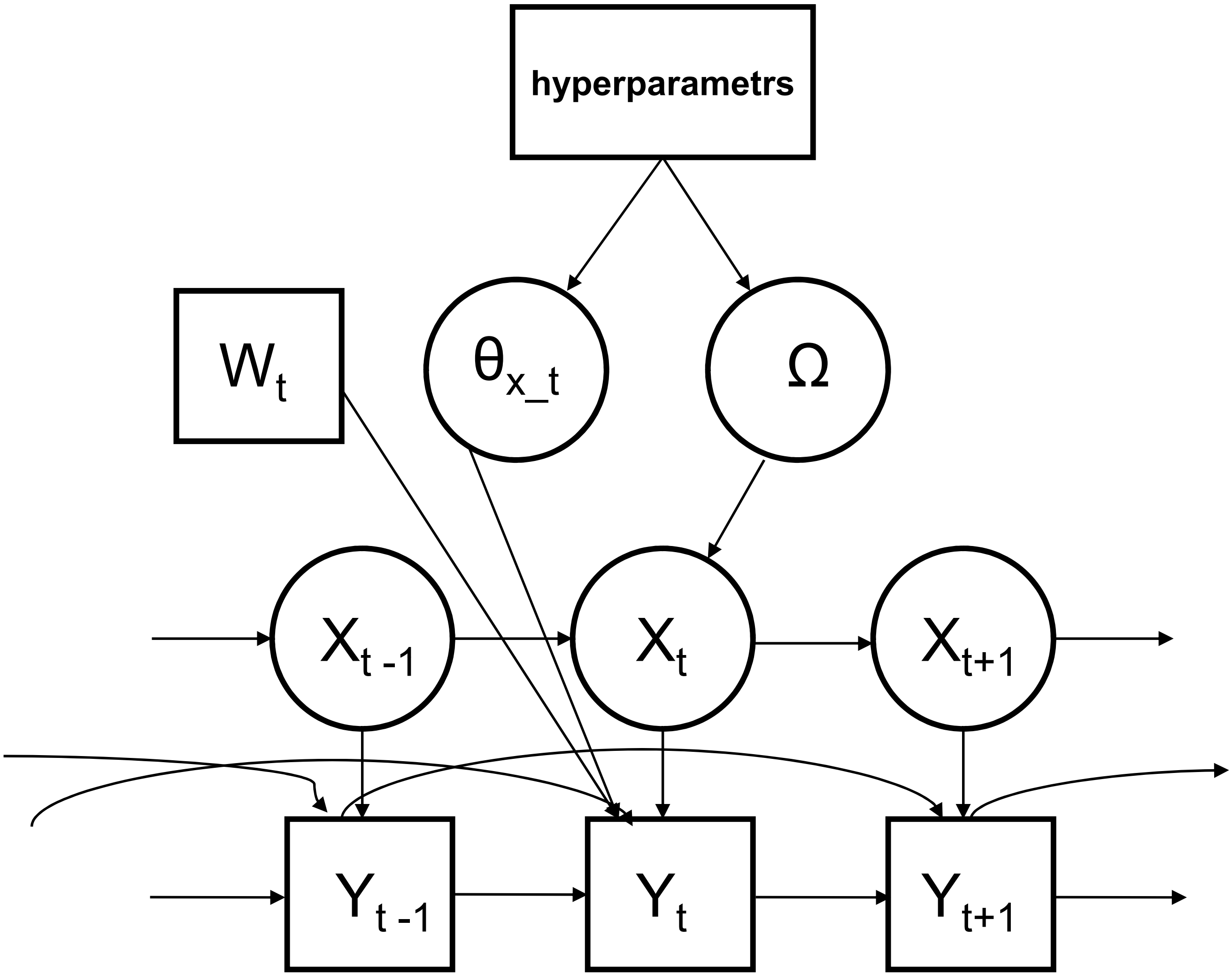
Non-homogeneous hidden Markov models



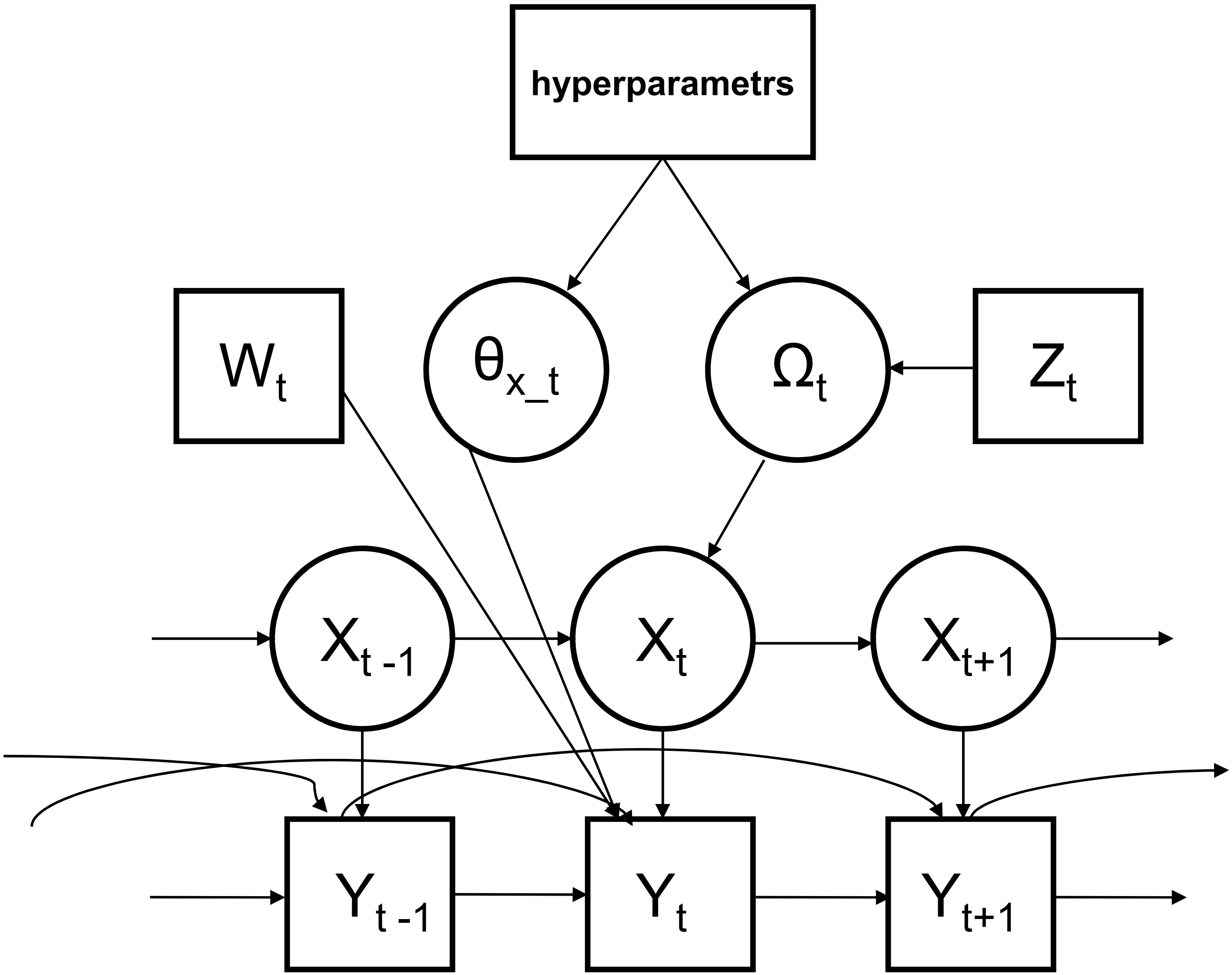
Markov switching autoregressive models



Markov switching autoregressive models + covariates



Markov switching autoregressive models + covariates



Variable selection methods

Stochastic Search Variable Selection (George and McCullogh, 1993)

SSVS

Kuo and Mallick’s method (Kuo and Mallick, 1998)

KM

Gibbs Variable Selection (Dellaportas, Forster, Ntzoufras, 2000)

GVS

Metropolized-Kuo-MallicK (Paroli and Spezia, 2008)

MKMK

Variable selection methods

$w_h = (w_{1,h}, \dots, w_{t,h}, \dots, w_{T,h})$ $\qquad h=1,\dots,q$

$\gamma_h = 1 \Rightarrow w_h$ included

$\gamma_h = 0 \Rightarrow w_h$ excluded

There are 2^q possible models to select

The best model is identified by its highest posterior probability, that is the subset of covariates corresponding to the vector $(\gamma_1,\dots, \gamma_h,\dots,\gamma_q)$ with the highest frequence of appearence in the MCMC sample

(Gaussian) Hidden Markov models

$(y_t | x_t = i) = \mu_i + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1}) \qquad (i = 1,\ldots,m)$

$P(x_t = i \mid x_{t-1} = k) = \omega_{k,i}$

$(y_t | x_t = i) \sim \mathcal{N}\left(\mu_i + e_t; \lambda_i^{-1}\right)$

$y_t \sim \sum_{i=1}^{m} \omega_{j,i} \mathcal{N}\left(\mu_i; \lambda_i^{-1}\right)$

Non-Homogeneous Hidden Markov models

$\Omega_t = \left[\omega^t_{j,i}\right]$ $\quad$ $\mathrm{logit}(\omega^t_{j,i}) = \ln(\omega^t_{j,i}/\omega^t_{i,i}) = z_t'\alpha_{j,i}$ $\quad$ $z_t = (1, z_{t,1}, \ldots, z_{t,q})'$

$\alpha_{j,i} = (\alpha_{0(j,i)}, \alpha_{1(j,i)}, \ldots, \alpha_{q(j,i)})'$ if $i \neq j$

$\alpha_{i,j} = 0_{(q)}$ if $i = j$

$$\omega^t_{j,i} = \frac{\exp(z_t'\alpha_{j,i})}{1 + \sum_{i=1}^{m} \exp(z_t'\alpha_{j,i})}$$

Stochastic Search Variable Selection

$(y_t \mid x_t = i) = \mu_i + e_t$

$$\omega^t_{j,i} = \frac{\exp(z_t'\alpha_{j,i})}{1 + \sum_{i=1}^{m} \exp(z_t'\alpha_{j,i})}$$

$\gamma = (\gamma_1', \ldots, \gamma_j', \ldots, \gamma_m')'$;

$\gamma_{(j)} = (1, \gamma_{1(j)}, \ldots, \gamma_{q(j)})'$

$\gamma_{k(j)} = 1 \Rightarrow z_{t-1,k}$ included, if $x_{t-1} = j$

$\gamma_{k(j)} = 0 \Rightarrow z_{t-1,k}$ excluded, if $x_{t-1} = j$

$\mu_i \sim \mathcal{N}(\bullet; \bullet)$ $\lambda_i \sim \mathcal{G}(\bullet; \bullet)$

$\alpha_{j,i} \mid \gamma_j \sim \mathcal{N}_{q+1}(0; D_{\gamma(j)})$

$D_{\gamma(j)} = \text{diag}[1, (\delta_{1(j)}\tau_{1(j)})^2, \ldots, \delta_{k(j)}\tau_{k(j)})^2, \ldots, (\delta_{q(j)}\tau_{q(j)})^2]$

with $\delta_{k(j)} = c_{k(j)}$ if $\gamma_{k(j)} = 1$ and $\delta_{k(j)} = 0$ and $\gamma_{k(j)} = 1$; $c_{k(j)}$ and $\tau_{k(j)}$ fixed

$\gamma_{k(j)} \sim \mathcal{B}e(0.5)$

Kuo and Mallick's method

$(y_t | x_t = i) = \mu_i + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1}) \qquad (i = 1,\ldots,m)$

$$\omega^t_{j,i} = \frac{\exp(z_t' \operatorname{diag}[\gamma_j] \alpha_{j,i})}{1 + \sum_{i=1}^{m} \exp(z_t' \operatorname{diag}[\gamma_j] \alpha_{j,i})}$$

$\mu_i \sim \mathcal{N}(\bullet; \bullet)$

$\lambda_i \sim \mathcal{G}(\bullet; \bullet)$

$\alpha_{j,i} \sim \mathcal{N}_{q+1}(\bullet; \bullet)$

$\gamma_{k(j)} \sim \mathcal{B}e(0.5)$

SSVS + KM

$(y_t | x_t = i) + \mu_i + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1}) \qquad (i = 1,\dots,m)$

$$\omega^t_{j,i} = \frac{\exp(z_t' \operatorname{diag}[\gamma_j]\, \alpha_{j,i})}{1 + \sum_{i=1}^{m} \exp(z_t' \operatorname{diag}[\gamma_j]\, \alpha_{j,i})}$$

$\mu_i \sim \mathcal{N}(\bullet; \bullet) \qquad \lambda_i \sim \mathcal{G}(\bullet; \bullet)$

$\alpha_{j,i} | \gamma_j \sim \mathcal{N}_{q+1}(0; D_{\gamma(j)})$

$$D_{\gamma(j)} = \operatorname{diag}[1, (\delta_{1(j)}\tau_{1(j)})^2, \dots, \delta_{k(j)}\tau_{k(j)})^2, \dots, (\delta_{q(j)}\tau_{q(j)})^2]$$

with $\delta_{k(j)} = c_{k(j)}$ if $\gamma_{k(j)} = 1$ and $\delta_{k(j)} = 0$ and $\gamma_{k(j)} = 1$; $c_{k(j)}$ and $\tau_{k(j)}$ fixed

$\gamma_{k(j)} \sim \mathcal{Be}(0.5)$

Metropolized-Kuo-MallicK

$(y_t | x_t = i) = \mu_i + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1}) \qquad (i = 1,\ldots,m)$

$$\omega^t_{j,i} = \frac{\exp(z_t' \operatorname{diag}[\gamma_j]\, \alpha_{j,i})}{1 + \sum_{i=1}^{m} \exp(z_t' \operatorname{diag}[\gamma_j]\, \alpha_{j,i})}$$

$\mu_i \sim \mathcal{N}(\bullet; \bullet)$

$\lambda_i \sim \mathcal{G}(\bullet; \bullet)$

$\alpha_{j,i} \sim \mathcal{N}_{q+1}(\bullet; \bullet)$

$\gamma_{k(j)} \sim \mathcal{Be}(0.5)$

Simulations

n = 500; q = 5; m = 2; 3

corr(Z_2; Z_4) = corr(Z_1; Z_5) = 0; 0.3; 0.7; 0.9
other corr = random; |random| $\leq$ corr(Z_2; Z_4)

ex: n = 500; q = 5; m = 3; corr(Z_2; Z_4) = corr(Z_1; Z_5) = 0.7

	SSVS	KM	GVS	MKMK
state 1 (Z_4; Z_5)	.04*	.07*	.13	.64
state 2 (Z_2; Z_3 ; Z_5)	.04*	.08*	.12	.44
state 3 (Z_2; Z_5)	.03*	.05*	.10	.70

Markov switching autoregressive models

$$(y_t | x_t = i) = \mu_i + \sum_{\tau=1}^{p} \varphi_{\tau(i)}\, y_{t-\tau} + \sum_{h=1}^{q} \theta_{h(i)}\, w_{t,h} + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1})$$

$$(i = 1, \ldots, m)$$

$$P(x_t = i \mid x_{t-1} = k) = \omega_{k,i}$$

$$(y_t | x_t = i) \sim \mathcal{N}\left(\mu_i + \sum_{\tau=1}^{p} \varphi_{\tau(i)}\, y_{t-\tau};+ \sum_{h=1}^{q} \theta_{h(i)}\ \ w_{t,h};\ \lambda_i^{-1}\right)$$

$$y_t \sim \sum_{i=1}^{m} \omega_{j,i}\, \mathcal{N}\left(\mu_i + \sum_{\tau=1}^{p} \varphi_{\tau(i)}\, y_{t-\tau};+ \sum_{h=1}^{q} \theta_{h(i)}\ \ z_{t,h};\ \lambda_i^{-1}\right)$$

Stochastic Search Variable Selection

$$(y_t | x_t = i) = \mu_i + \sum_{\tau=1}^{p} \varphi_{\tau(i)} y_{t-\tau} + \sum_{h=1}^{q} \theta_{h(i)} w_{t,h} + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1})$$

$\gamma_i = (\gamma_{1(i)}, \ldots, \gamma_{h(i)}, \ldots, \gamma_{q(i)})$ for any i

$\mu_i \sim \mathcal{N}(\bullet; \bullet)$

$\lambda_i \sim \mathcal{G}(\bullet; \bullet)$

$\theta_{h(i)} | \boldsymbol{\gamma_{h(i)}} \sim [\gamma_{h(i)} * \mathcal{N}(0; c_{h(i)}^2 \tau_{h(i)}^2) + (1 - \gamma_{h(i)}) * \mathcal{N}(0; \tau_{h(i)}^2)]$

$c_{h(i)}$ and $\tau_{h(i)}$ fixed

$\gamma_{h(i)} \sim \mathcal{B}e(0.5)$

Kuo and Mallick's method

$$(y_t | x_t = i) = \mu_i + \sum_{\tau=1}^{p} \varphi_{\tau(i)} y_{t-\tau} + \sum_{h=1}^{q} \theta_{h(i)} \boldsymbol{\gamma_{h(i)}} w_{t,h} + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1})$$

$\gamma_i = (\gamma_{1(i)}, \ldots, \gamma_{h(i)}, \ldots, \gamma_{q(i)})$ for any i

$\mu_i \sim \mathcal{N}(\bullet; \bullet)$

$\lambda_i \sim \mathcal{G}(\bullet; \bullet)$

$\theta_{h(i)} \sim \mathcal{N}(\bullet; \bullet)$

$\gamma_{h(i)} \sim \mathcal{Be}(0.5)$

Gibbs Variable Selection

$$(y_t | x_t = i) = \mu_i + \sum_{\tau=1}^{p} \varphi_{\tau(i)}\, y_{t-\tau} + \sum_{h=1}^{q} \theta_{h(i)}\, \boldsymbol{\gamma_{h(i)}}\, w_{t,h} + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1})$$

$$\gamma_i = (\gamma_{1(i)}, \ldots, \gamma_{h(i)}, \ldots, \gamma_{q(i)}) \qquad \text{for any } i$$

$$\mu_i \sim \mathcal{N}(\bullet; \bullet)$$

$$\lambda_i \sim \mathcal{G}(\bullet; \bullet)$$

$$\theta_{h(i)} | \boldsymbol{\gamma_{h(i)}} \sim [\gamma_{h(i)} * \mathcal{N}(0; c_{h(i)}^2 \tau_{h(i)}^2) + (1 - \gamma_{h(i)}) * \mathcal{N}(0; \tau_{h(i)}^2)]$$

$c_{h(i)}$ and $\tau_{h(i)}$ fixed

$$\gamma_{h(i)} \sim \mathcal{Be}(0.5)$$

Metropolized-Kuo-MallicK

$$(y_t | x_t = i) = \mu_i + \sum_{\tau=1}^{p} \varphi_{\tau(i)}\, y_{t-\tau} + \sum_{h=1}^{q} \theta_{h(i)}\, \boldsymbol{\gamma_{h(i)}}\, w_{t,h} + e_t \qquad e_t \sim \mathcal{N}(0; \lambda_i^{-1})$$

$\gamma_i = (\gamma_{1(i)}, \ldots, \gamma_{h(i)}, \ldots, \gamma_{q(i)})$ for any i

$\mu_i \sim \mathcal{N}(\bullet; \bullet)$

$\lambda_i \sim \mathcal{G}(\bullet; \bullet)$

$\theta_{h(i)} \sim \mathcal{N}(\bullet; \bullet)$

$\gamma_{h(i)} \sim \mathcal{B}e(0.5)$

Simulations

$n = 500$ $q = 5$ $m = 2; 3$ $p = 1; 2; 3$

$corr(Z_2; Z_4) = corr(Z_1; Z_5) = 0; 0.3; 0.7; 0.9$

other corr = random; $|random| \leq corr(Z_2; Z_4)$

ex: $n = 500; q = 5; m = 3; p = 3; corr(Z_2; Z_4) = corr(Z_1; Z_5) = 0.7$

	SSVS	KM	GVS	MKMK
state 1 ($Z_4; Z_5$)	.26	.58	.81	.61
state 2 ($Z_2; Z_3 ; Z_5$)	.07	.41	.82	.62
state 3 ($Z_2; Z_5$)	.18	.61	.84	.60

Three applications

Bernoulli non-homogeneous hidden Markov model for mapping species distribution in a river

Non-homogeneous Markov switching autoregressive models + covariates for the analysis of water and air quality

Freshwater pearl mussels

Presence or absence of freshwater pearl mussels in the River Dee, Scotland

http://www.snh.gov.uk/

1213 sections *t*

In each section:

$y_t=1 \Rightarrow$ mussels observed

$y_t=0 \Rightarrow$ mussels not observed

$x_t=1 \Rightarrow$ presence of mussels

$x_t=0 \Rightarrow$ absence mussels not observed

42 missing values

Freshwater pearl mussels

Bernoulli non-homogeneous hidden Markov model

$m = 2$

34 covariates

Ω_t

MKMK for variable selection

$\omega_{0,1} = \omega(\text{-bridges-dredging-wwtw})$

$\omega_{1,0} = \omega(\text{+tributaries+wwtw})$

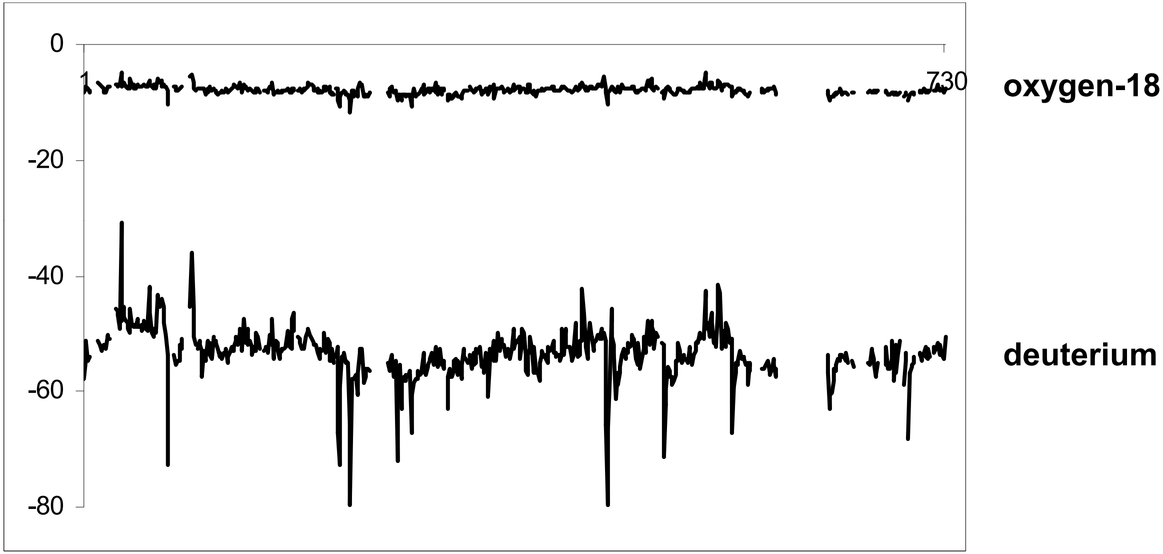
Isotopes

Daily concentrations of oxygen-18 and deuterium, sampled in the Wemyss catchment in eastern Scotland

Two years of daily data (T = 730)

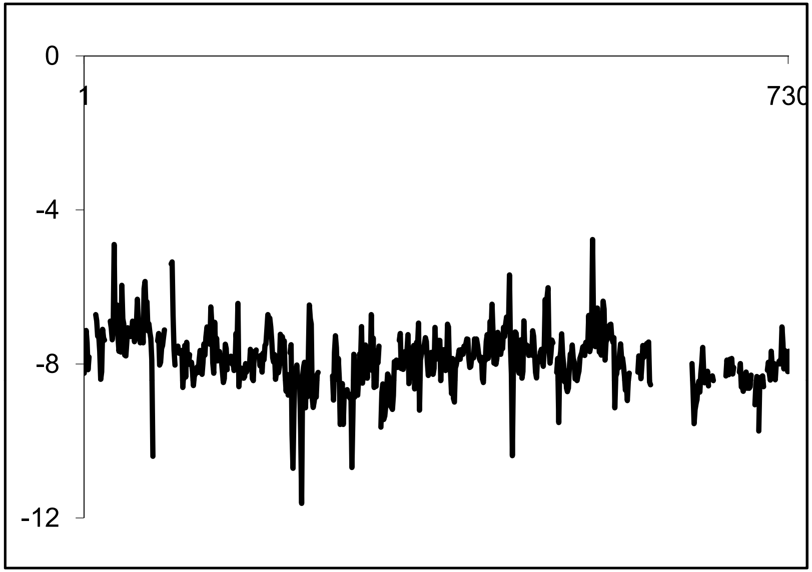
127 missing values (17%)

9 covariates

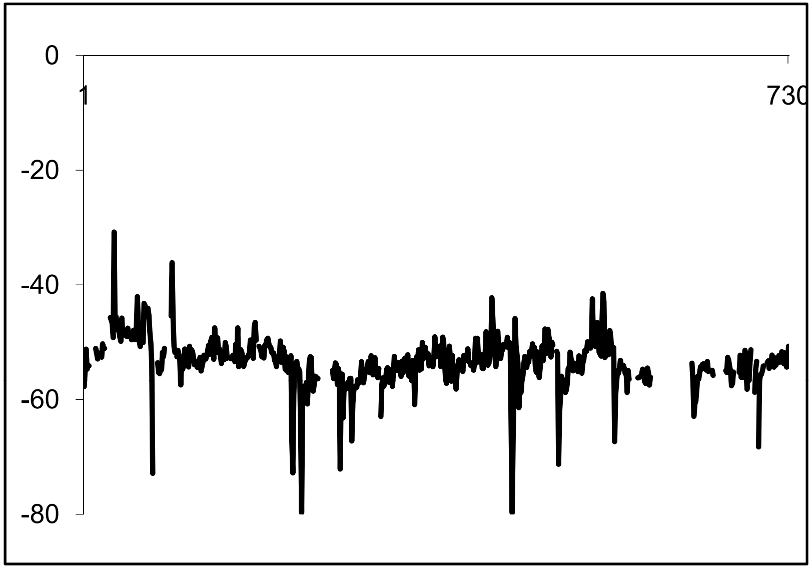


Isotopes

Non-homogeneous Markov switching autoregressive models + covariates + yearly periodic component



oxygen-18



deuterium

Model choice:

Marginal Likelihoods From the Metropolis-Hastings output (Chib and Jeliazkov, 2001)

oxygen-18: $m = 2$
$p = 1$

deuterium: $m = 2$
$p = 2$

Variable selection:
MKMK method (contemporary covariates)

oxygen-18

$\{y_t\}$ - state 1: T, P_18O
state 2: T

$\{x_t\}$ - state 1: P_18O
state 2: Tu, P_18O

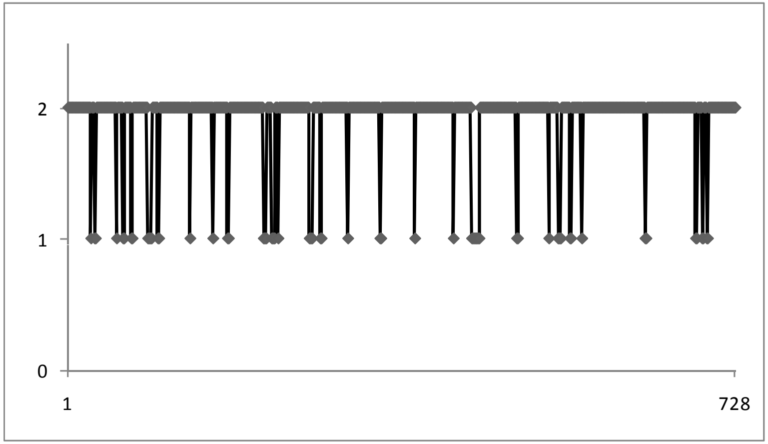
deuterium

$\{y_t\}$ - state 1: Tu, T, P_D
state 2: T

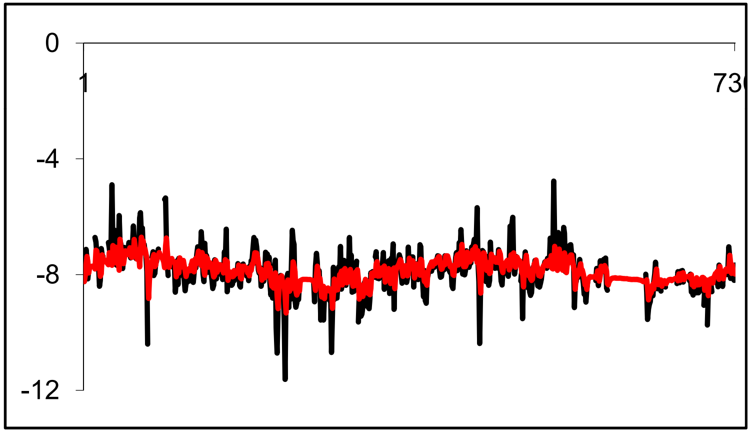
$\{x_t\}$ - state 1: P
state 2: P_D

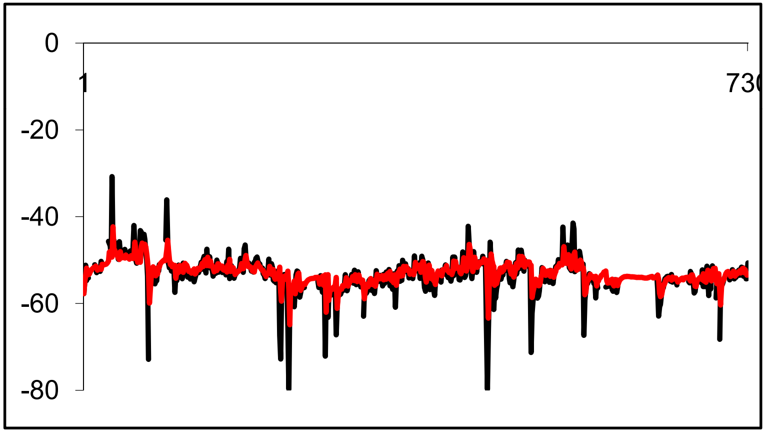
Parameter estimation:

oxygen-18

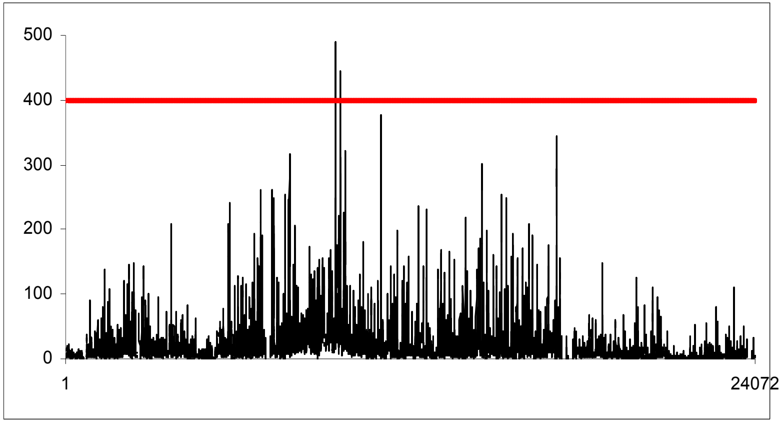


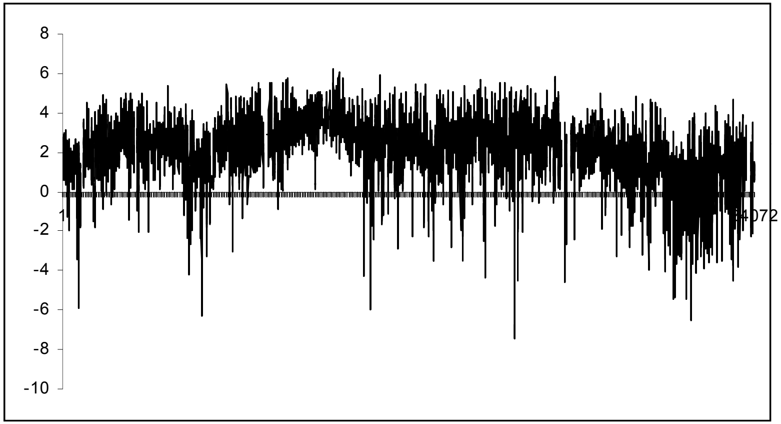
deuterium





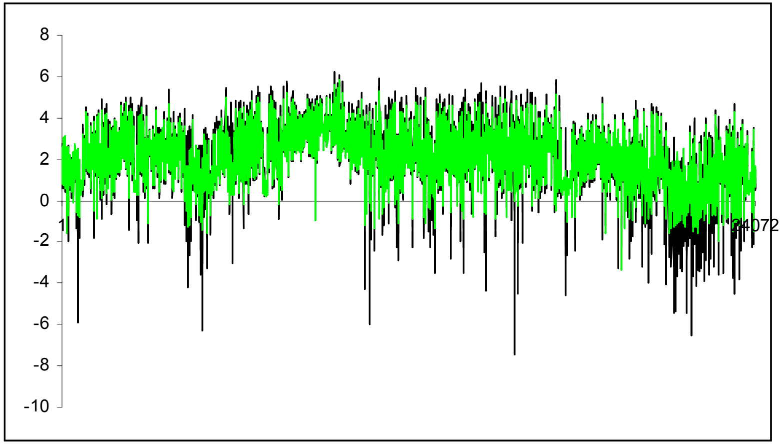
SO2

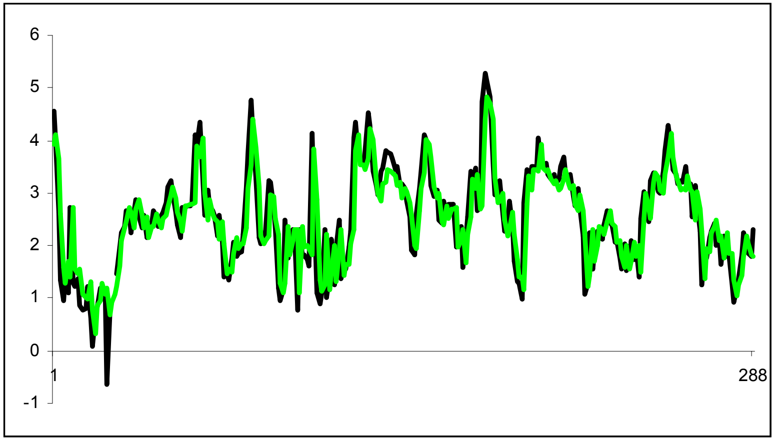




hourly mean concentrations of sulphur dioxide (SO_2), measured in μg/m^3, recorded on the Isle of Giudecca (lagoon of Venice), from 1.1.2001 to 30.9.2003 (24072 observations)

SO2





January 1st, 2001 / September 30th, 2003

September 26th / October 7th, 2001

BLACK observed values

GREEN fitted values

SO2

w state 1: Wind Speed

state 2: Temperature

state 3: Wind Speed

state 4: Wind Speed, Temperature

z state 1: Temperature

state 2: Temperature

state 3: Solar Radiation

state 4: Temperature, Atmospheric Pressure

Summary

By simulation studies we compared the performance of SSVS, KM, GVS, MKMK methods when applied to MSARMs and NHHMMs

2 applications of MKMK to NHMSARMs

Application of MKMK to a Bernoulli NHHMM

Current and future work

Negative Binomial non-homogeneous hidden Markov models and freshwater pearl mussels counts in River Dee

Spatial hidden Markov models and birds distributions in South Africa

Multivariate non-homogeneous Markov switching models and isotopes